Intern Emerg Med (2018) 13:449–451
https://doi.org/10.1007/s11739-017-1763-0

CE - CLINICAL NOTES

Understanding and improving decisions in clinical medicine (III): towards cognitively informed clinical thinking

Vincenzo Crupi[1] · Fabrizio Elia[2] · Franco Aprà[2]

Received: 29 August 2017 / Accepted: 14 October 2017 / Published online: 23 October 2017

Clinical judgment under uncertainty

Clinical practice is often associated with judgment and with uncertainty, and rightly so. Since the logic of uncertainty is probability theory, understanding clinical judgment requires consideration of how clinicians assess probabilities. Suppose, for instance, that you are involved in a mammography screening program for the early detection of breast cancer. A 50-year-old woman with no symptoms has a positive test result. The pretest probability of breast cancer in her age group is 1%, and the sensitivity and specificity of the test are 80 and 90%, respectively (so the false positive rate is 10%). In light of her positive mammography, what is the probability that your patient actually has breast cancer?

A problem of this kind represents a crucial fragment of diagnostic reasoning. It involves an uncommon and serious disease, a useful but imperfect item of evidence, and it demands a judgment about the former on the basis of the latter (as a key example in emergency medicine, one can think of the use of troponin for myocardial infarction). Arguably, for more nuanced and challenging clinical problems to be solved effectively, one should be able to handle something like the mammography problem correctly. Thus, it was striking and unsettling when researchers first found out that physicians' intuitive assessments were largely off the mark. As early as 1982, David Eddy (credited for having introduced the expression "evidence-based medicine"), reported that physicians' estimates of breast cancer turned out to be mistaken by up to one order of magnitude: close to 80%, the correct answer being about 7.5% (see below) [1].

The cognitive mechanism leading to so high an estimate is a variant of so-called *representativeness heuristic*. Many clinicians derive their inflated judgment of the positive predictive value—P(cancer|test+)—from the consideration that a positive mammogram is a very representative (or typical) feature of a woman with cancer, as indicated by the high sensitivity of the test, or by the sizable likelihood ratio of a positive result (80%/10% = 8).

Format can be of substance

The mistake of intuitively overestimating the positive predictive value of a valid test for an uncommon disease has proven to be widespread regardless of the specific clinical content. The obvious question is then how can this pitfall be avoided? To address this point, consider a variant of the mammography problem. Imagine 1000 women like your patient. Ten out of 1000 actually have a breast cancer (that is, 1%). Of those ten women with breast cancer, eight will have a positive mammography upon screening. Among the 990 women who do not have breast cancer, 99 will still have a (falsely) positive mammography. So how many of the women with a positive test result really have breast cancer? ____ out of ____

In a 2000 *Science* paper, Gerd Gigerenzer and collaborators investigated advanced medical students' judgment with this kind of formulation, which they called "natural

✉ Vincenzo Crupi
vincenzo.crupi@unito.it

Fabrizio Elia
fabrizioelia@libero.it

Franco Aprà
franco.apra@aslto2.piemonte.it

1 Department of Philosophy and Education, Center for Logic, Language, and Cognition, University of Turin, Turin, Italy

2 High Dependency Unit, San Giovanni Bosco Hospital, Turin, Italy

frequency format". Employing four different clinical scenarios, they found remarkable improvement. On average, rephrasing the problems increased the proportion of correct responses from about 20% to the majority of participants [2]. In the mammography case, for instance, most clinicians were able to realize that eight women with cancer plus 99 without cancer from the initial sample are expected to have a positive result, so that the positive predictive value corresponds to 8 out of 107 (8 + 99), that is, just about 7.5% probability.[1] In this line of reasoning, the low base-rate or prevalence plays its due role, because the substantial number of expected false positive results (99) arises as a small proportion of a large group of people without the disease (10% of 990). Notably, the effect observed depended on the change in format alone: no additional intervention was involved.

Cognitive tools for improving reasoning

If human cognition was primarily driven by the application of optimal formal rules, a percentage vs. nested frequency representation would amount to an inconsequential variation in how the same statistical information is encoded: it should not matter. And yet, it turns out to have important effects for clinicians' thinking. The key achievement of the cognitive science of human reasoning applies once again: judgments and preferences are typically construed through—not just revealed by—the decision process. Accordingly, depending on factors of context and format, which favor specific cognitive mechanisms, they can be biased or inconsistent in systematic and predictable ways. The good news is that, for the same kind of reasons, acting on factors that might look inconsequential from a strictly logical perspective can affect human cognition positively, for instance by making the correct solution of a problem more transparent.

This case of study is a classic proof of concept of the cognitive science of clinical reasoning. With diagnostic tests, the standard language of probabilities (prevalence, sensitivity, specificity) is essential in the scientific literature, yet it has been known for decades to prompt systematical flaws in clinical reasoning. Rephrasing the mammography case and other similar problems in terms of nested frequencies is an efficient remedial strategy. It is simple to teach and learn. Moreover, it defuses a broadly useful pattern of heuristic reasoning (to wit, representativeness) in an appropriately selective way, that is, in a specific subset of judgment tasks where the risk of error is known in advance to be otherwise significant.

Recently, we presented the usual mammography problem once again to 78 relatively young physicians (mean age 34; 69% female). They were asked to select "the most accurate estimate" of the probability of breast cancer among four options: 1, 10, 50, and 80%. For the appropriate response ("10%"), performance was virtually at chance level (chosen by 27%). The modal response was "80%" (41% of the choices), and 18% of the participants chose "50%", so that the majority (59% overall) judged the probability of cancer to be one half or more. Apparently, the tools to avoid this mistake are not being taught and disseminated. Consider a popular textbook on *Medical Statistics Made Easy* [4], reprinted three times over the last 15 years. One section is devoted to the interpretation of test results, and the explanation given is formally neat. It is also complemented with a telling final alert which might puzzle the reader: "if you are still feeling confused, you are in good company". In fact, no mention is ever made of the specific kinds of reasoning flaws that are known to be prevalent in these problems, nor is the cognitive relevance of format explicitly addressed.

Conclusion

Nowadays, hardly anyone would question that clinical decisions should be made on the basis of scientific evidence. Few seem to have realized, however, that cognitive psychology, too, is one basic science for medical decision making [5]. Combining these two ideas remains an important task in contemporary medicine. Medical education should become fully evidence-based, including relevant evidence from cognitive science, and it should aim at making routine clinical thinking more cognitively informed, thus more vigilant against reasoning errors that are both predictable and avoidable.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Statement of human and animal rights This article does not contain any studies with human and animals performed by any of the authors.

Informed consent Informed consent was obtained from all participants included in this study.

[1] Meanwhile, other researchers have shown that the frequency terminology is not even necessary to foster accurate responses, as long as the right kind of nested structure is conveyed in the representation of the problem [3].

References

1. Eddy DM (1982) Probabilistic reasoning in clinical medicine: problems and opportunities. In: Kahneman D, Slovic P, Tversky A (eds) Judgment under uncertainty: heuristics and biases. Cambridge University Press, Cambridge, pp 249–267
2. Hoffrage U, Lindsey S, Hertwig R, Gigerenzer G (2000) Communicating statistical information. Science 290:2261–2262
3. Girotto V, Gonzalez M (2001) Solving probabilistic and statistical problems: a matter of information structure and question form. Cognition 78:247–276
4. Harris M, Taylor G (2014) Medical statistics made easy, 3rd edn. Taylor & Francis, London
5. Redelmeier DA, Ferris LE, Tu JV, Hux JE, Schull MJ (2001) Problems for clinical judgment: introducing cognitive psychology as one more basic science. CMAJ 164:358–360